

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2002, 13:38:49 ; Search time 13.36 Seconds
(without alignments)
1772.865 Million cell updates/sec

Title: US-09-405-504A-25

Perfect score: 3372

Sequence: 1 MRAPGAGAAVVVSLALLWLL.....HYLPNEAVYTRICSGAFAL 646

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3062	90.8	646	1	FATP_RAT
2	3054	90.6	646	1	FATP_MOUSE
3	1046.5	31.0	620	1	VLCS_HUMAN
4	1044.5	31.0	620	1	VLCS_MOUSE
5	1032.5	30.6	620	1	VLCS_MOUSE
6	699	20.7	623	1	FAT1_YEAST
7	390.5	11.6	522	1	CAIC_ECOLI
8	301.5	8.9	2560	1	PF52_BACSU
9	298.5	8.9	544	1	4CL1_PETCR
10	297.5	8.8	544	1	4CL1_PETCR
11	297.5	8.7	3587	1	SRF1_BACSU
12	290.5	8.6	547	1	4CL1_TOBAC
13	287.5	8.5	3587	1	SRF2_BACSU
14	281.5	8.3	561	1	4CL3_ARATH
15	281	8.3	563	1	4CL1_ORYSA
16	280	8.3	548	1	4CL1_ORYSA
17	278.5	8.2	545	1	4CL2_SOLTU
18	277.5	8.2	545	1	4CL1_SOLTU
19	272.5	8.1	553	1	4CL1_VANPL
20	268.5	8.0	542	1	4CL2_TOBAC
21	268	7.9	561	1	LCFA_ECOLI
22	259.5	7.7	661	1	ACSA_COPCI
23	256.5	7.6	537	1	4CL1_PINTA
24	256.5	7.6	651	1	ACSA_MYCTU
25	249	7.4	6486	1	TYCC_BACBR
26	245	7.3	561	1	4CL1_ARATH
27	245	7.3	562	1	4CL2_SOYBN
28	244.5	7.3	4451	1	GRSB_BACBR
29	241.5	7.2	3587	1	TYCB_BACBR
30	237.5	7.0	572	1	ACSA_BACSU
31	237	7.0	562	1	LCFA_HAEIN
32	235.5	7.0	556	1	4CL2_ARATH
33	235	7.0	628	1	PRPE_ECOLI

34	234	6.9	628	1	PRPE_SALTY
35	233	6.9	683	1	ACS2_YEAST
36	232	6.9	684	1	ACS2_KLULA
37	232	6.9	805	1	ACVS_STROL
38	230	6.8	670	1	ACSA_EMENT
39	230	6.8	5255	1	BACA_BACLI
40	228.5	6.8	6359	1	BACC_BACLI
41	227	6.7	543	1	FAT2_YEAST
42	226.5	6.7	675	1	ACSA_CANAL
43	226.5	6.7	2555	1	PPS1_BACSU
44	225.5	6.7	2561	1	PPS1_BACSU
45	224.5	6.7	669	1	ACSA_PENCH

ALIGNMENTS

RESULT	1
FATP_RAT	
ID	FATP_RAT
AC	STANDARD; PRT; 646 AA.
DT	01-NOV-1997 (Rel. 35, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	20-AUG-2001 (Rel. 40, Last annotation update)
DE	LONG-CHAIN FATTY ACID TRANSPORT PROTEIN (FATP).
GN	SLC27A1 OR FATP.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX	NCBI_TaxID=10116;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=98041635; PubMed=9375787;
RT	Schaap F.G., Hamers L., van der Vusse G.J., Glatz J.F.C.;
RA	"Molecular cloning of fatty acid-transport protein cDNA from rat.";
RL	Biochim. Biophys. Acta 1354:29-34(1997).
CC	-!- FUNCTION: INVOLVED IN TRANSLLOCATION OF LONG-CHAIN FATTY ACIDS
CC	ACROSS THE PLASMA MEMBRANE. MAY PLAY A PIVOTAL ROLE IN REGULATING
CC	AVAILABLE LONG-CHAIN FATTY ACID SUBSTRATES FROM EXOGENOUS SOURCES
CC	IN TISSUES UNDERGOING HIGH LEVELS OF BETA-OXIDATION OR
CC	TRIGLYCERIDE SYNTHESIS.
CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PLASMA MEMBRANE.
CC	-!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC	FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U89529; AAC53424.1; ..
InterPro; IPR000873; AMP-bind.
Pfam; PF00501; AMP-binding; 1.
PROSITE; PS00455; AMP_BINDING; 1.
Glycoprotein; Lipid transport; Transmembrane.
TRANSMEM 14 34
TRANSMEM 115 135
TRANSMEM 140 160
TRANSMEM 169 189
TRANSMEM 293 313
TRANSMEM 537 557
TRANSMEM 330 330
CARBOHYD 393 393
CARBOHYD 518 518
SEQUENCE 646 AA; 71283 MW; C450CF174CC2EB87 CRC64;

Query Match 90.8%; Score 3062; DB 1; Length 646;
Best Local Similarity 89.5%; Pred. No. 2e-231;
Matches 578; Conservative 30; Mismatches 38; Indels 0; Gaps 0;

```
QY 1 MRAPGAGAAVSVSLALLWLLGLPWTWSAAALGVYVGGWRFLRIVCKTARRDLFGLSV 60
DB 1 MPTGAGTASVASGLLWLLGLPWTWSAAAFGVYVGGWRFLRIVCKTARRDLFGLSV 60
QY 61 LIRVLELRHRRAGHTTIPRIFOAVVQRPRLALVDAGTGCWTFQAOLDAYSNAVANLF 120
DB 61 LIRVLELRHRRAGHTTIPRIFOAVVQRPRLALVDAGTGCWTFQAOLDAYSNAVANLF 120
QY 121 ROLGFPAGDVVAIFLEGPRFVGLWGLAKAGMEALLNVLNRLREPLAFCLGTSAAKALI 180
DB 121 ROLGFPAGDVVAIFLEGPRFVGLWGLAKAGMEALLNVLNRLREPLAFCLGTSAAKALI 180
QY 181 FCGEMAAVAEVSCHLGSLLKFCSDGLGPEGLPDTLLDPLKKEASTAPLAQIPSKGM 240
DB 181 YGEMAAVAEVSCHLGSLLKFCSDGLGPEGLPDTLLDPLKKEASTAPLAQIPSKGM 240
QY 241 DDRLFYIYTSGLTGLPKAAIIVVHSRYRMAAFGHSHYRMAADVLYDCLPLYHSAGNIIG 300
DB 241 DDRLFYIYTSGLTGLPKAAIIVVHSRYRMAAFGHSHYRMAADVLYDCLPLYHSAGNIIG 300
QY 301 VGOCIIYGLTVLWLRKFSASRFDWDCIKYNTVQYIGEICRYLLKOPVREARRHRVRL 360
DB 301 VGOCIIYGLTVLWLRKFSASRFDWDCIKYNTVQYIGEICRYLLKOPVREARRHRVRL 360
QY 361 AVGNGLRPAIWEETFERFVGRQIGEFYGAECNCSIANMDGKVGSCGFNSRLPHVYPYR 420
DB 361 AVGNGLRPAIWEETFERFVGRQIGEFYGAECNCSIANMDGKVGSCGFNSRLPHVYPYR 420
QY 421 LKVNEDTMEPLRQAGLIPCQAGEPGLLVQINQDPLRRFDGYVSDSNKTKIAHSV 480
DB 421 LKVNEDTMEPLRQAGLIPCQAGEPGLLVQINQDPLRRFDGYVSDSNKTKIAHSV 480
QY 481 FSKGDSAYLSGDLVYMDLGYMYFRDRSGDTFRKRGENVSTTEVEGLSRLLGOTDVAVY 540
DB 481 FRKGSAYLSGDLVYMDLGYMYFRDRSGDTFRKRGENVSTTEVEGLSRLLGOTDVAVY 540
QY 541 GVAVPVGESKAGMAAVADPHSLDPAIYQELQKVLAPYRPIFLRLPQVDVTTGTFKIQ 600
DB 541 GVAVPVGESKAGMAAIADPHNQLDPSMYQELQKVLASYAQPIFLRLPQVDVTTGTFKIQ 600
QY 601 KTRLOREGDPQTSRDLFFLDLKGCHYLPLNEAVYTRICSGAFAL 646
DB 601 KTRLOREGDPQTSRDLFFLDLKGCHYLPLNEAVYTRICSGAFAL 646
RESULT 2
FATP_MOUSE STANDARD: PRT: 646 AA.
AC Q60714;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LONG-CHAIN FATTY ACID TRANSPORT PROTEIN (FATP).
GN SLC27A1 OR FATP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS;
RX MEDLINE=95042740; PubMed=7954810;
RA Schaffer J.E., Lodish H.F.;
RT "Expression cloning and characterization of a novel adipocyte long
RT chain fatty acid transport protein.";
RL Cell 79:427-436(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98438516; PubMed=9765271;
RA Hui T.Y., Frohnert B.I., Smith A.J., Schaffer J.E., Bernlohr D.A.;
RT "Characterization of the murine fatty acid transport protein gene and
RT its insulin response sequence.";
```

```
RL J. Biol. Chem. 273:27420-27429(1998).
CC -!- FUNCTION: INVOLVED IN TRANSLLOCATION OF LONG-CHAIN FATTY ACIDS
CC ACROSS THE PLASMA MEMBRANE. MAY PLAY A PIVOTAL ROLE IN REGULATING
CC AVAILABLE LONG-CHAIN FATTY ACID SUBSTRATES FROM EXOGENOUS SOURCES
CC IN TISSUES UNDERGOING HIGH LEVELS OF BETA-OXIDATION OR
CC TRIGLYCERIDE SYNTHESIS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PLASMA MEMBRANE.
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN SKELETAL MUSCLE, HEART
CC AND FAT. LOWER LEVELS IN BRAIN, KIDNEY, LUNG AND LIVER. NO
CC EXPRESSION IN SPLEEN OR INTESTINE.
CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: UL5976; AAC71060.1; .
CC EMBL: AF023258; AAC69640.1; .
CC EMBL: AF023256; AAC69640.1; JOINED.
CC EMBL: AF023257; AAC69640.1; JOINED.
CC MGD: MGI:1347098; SLC27A1.
CC InterPro: IPR000873; AMP-bind.
CC Pfam: PF00501; AMP-binding; 1.
CC PROSITE: PS00455; AMP-BINDING; 1.
CC Glycoprotein; Lipid transport; Transmembrane.
KW TRANSMEM 14 34 POTENTIAL.
FT TRANSMEM 115 135 POTENTIAL.
FT TRANSMEM 140 160 POTENTIAL.
FT TRANSMEM 169 189 POTENTIAL.
FT TRANSMEM 293 313 POTENTIAL.
FT TRANSMEM 537 557 POTENTIAL.
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 646 AA; 71276 MW; 910B92BA8D985B4C CRC64;
Query Match 90.6%; Score 3054; DB 1; Length 646;
Best Local Similarity 89.5%; Pred. No. 8.6e-231;
Matches 578; Conservative 30; Mismatches 38; Indels 0; Gaps 0;
QY 1 MRAPGAGAAVSVSLALLWLLGLPWTWSAAALGVYVGGWRFLRIVCKTARRDLFGLSV 60
DB 1 MRAPGAGTASVASGLLWLLGLPWTWSAAAFGVYVGGWRFLRIVCKTARRDLFGLSV 60
QY 61 LIRVLELRHRRAGHTTIPRIFOAVVQRPRLALVDAGTGCWTFQAOLDAYSNAVANLF 120
DB 61 LIRVLELRHRRAGHTTIPRIFOAVVQRPRLALVDAGTGCWTFQAOLDAYSNAVANLF 120
QY 121 ROLGFPAGDVVAIFLEGPRFVGLWGLAKAGMEALLNVLNRLREPLAFCLGTSAAKALI 180
DB 121 ROLGFPAGDVVAIFLEGPRFVGLWGLAKAGMEALLNVLNRLREPLAFCLGTSAAKALI 180
QY 181 FCGEMAAVAEVSCHLGSLLKFCSDGLGPEGLPDTLLDPLKKEASTAPLAQIPSKGM 240
DB 181 YGEMAAVAEVSCHLGSLLKFCSDGLGPEGLPDTLLDPLKKEASTAPLAQIPSKGM 240
QY 241 DDRLFYIYTSGLTGLPKAAIIVVHSRYRMAAFGHSHYRMAADVLYDCLPLYHSAGNIIG 300
DB 241 DDRLFYIYTSGLTGLPKAAIIVVHSRYRMAAFGHSHYRMAADVLYDCLPLYHSAGNIIG 300
QY 301 VGOCIIYGLTVLWLRKFSASRFDWDCIKYNTVQYIGEICRYLLKOPVREARRHRVRL 360
DB 301 VGOCIIYGLTVLWLRKFSASRFDWDCIKYNTVQYIGEICRYLLKOPVREARRHRVRL 360
QY 361 AVGNGLRPAIWEETFERFVGRQIGEFYGAECNCSIANMDGKVGSCGFNSRLPHVYPYR 420
DB 361 AVGNGLRPAIWEETFERFVGRQIGEFYGAECNCSIANMDGKVGSCGFNSRLPHVYPYR 420
```


This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; D88308; BAA23644.1; -.
DB EMBL; AF096290; AAC64973.1; -.
DR MIM; 603247; -.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
KW Ligase; Fatty acid metabolism; Peroxisome; Transmembrane.
FT TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 107 127 POTENTIAL.
FT TRANSMEM 262 282 POTENTIAL.
SQ SEQUENCE 620 AA; 70312 MW; BFD33E1BB67B3DF8 CRC64;

Query Match 31.0%; Score 1044.5; DB 1; Length 620;
Best Local Similarity 40.1%; Pred. No. 7.1e-74;
Matches 254; Conservative 87; Mismatches 255; Indels 37; Gaps 15;

QY 34 VYVSGGWRFLRIVCKTARRDLF-GLSVLIRVLELRR-----HORAGHTIPRIFOAVVQ 87
DB 5 IYTVLAGLLFLPLNLNCCPFYFDIGYFLKAAVGRVRSYQGRRTILRAFLPKAR 64
QY 88 ROPERLALVDAGTGCWTFQAOLDAYSNAVAN-LFRQLGFAPGDVVAIFLEGRPEFVGLW 146
DB 65 QPHPKFLU--FRDETLTYAQRDRSNQVARALHDHGLRGDCVALLMGNEPAYVWLW 122
QY 147 GLAKAGMAEALLNVLNREPLAFCLGTSKAKALIPFGGEMVAEVSGLKGS--LTKFC 204
DB 123 GLVKLGCMACLNINRAKSLHLCPCCGAKVLLVSPQLQAAVEILPSLKDDVSIVYV 182
QY 205 SGDLGPEGILPDLTHLLDPLLEKASTAPLAQIPSKGMDORLF-----YIYTSCTTGLPKAA 259
DB 183 SRTSNTDGI--DSFL--DKVDEVSTPE--IPESWRSEVTFSTPALXIYTSCTTGLPKAA 235
QY 260 IYVHSRYRMAAFG--HHAYRMAQADVLCDPLPHYSAGNLIIGVGCGLIYGLVFWLRRKK 316
DB 236 MITHORIW---YGTGLTFVSGLKADDDIYITLFPYHSAALLIGHGCVAGATLALTK 291
QY 317 FSASRFDWDCIKYNTVQYIGICRYLLKQPVREAERHVRVLAAGNGLRPAIWEFE 376
DB 292 FSASQFWDCCRKYNTVQYIGICRYLLKQPVREAERHVRVLAAGNGLRPAIWEFE 351
QY 377 REGVRIQIGEFYCATCNSIANMDGKVGSCGNSRILPHVPIRLVKVNETMELLRDAQ 436
DB 352 REGDCIYEFYAATGEGNIGFMNARKVAGVRNLYLQKKIITYDLIKYDVEKDEPVRDEN 411
QY 437 GLICPCQAGEPGLLVQINQDPLRRFGYVSESATSKIAHSVFSKGSAYLSGDVIV 495
DB 412 GYCVVRPKEVGLLVCKITQLTP---FNGYAGAKAQTEKKLRDVFKKGDLVFNSGDLLM 468
QY 496 MDELGYMFRDRSGDTFRWRGENSTTEVEGVLKLLGTDVAVYGVAVPGVEGKAGNAA 555
DB 469 VDHENFIYFHDVRGDTFRWKGENVATTEVADTVGLVDFVQEVNIVYGVHVPDHEGRIGMAS 528
QY 556 V--ADPHSLDDPNALYQELQKVLAPYARPIFLRLLPQVDVTGTFKIQKTRLOREGEFDPQ 613
DB 529 IWKENHE--FDCKKLFQHIADYLPYARPRFLRIQDTITGTFKHKRWTLVEEGFNPAV 587
QY 614 TSDRLFFDLKQCHYLPLNEAVYTRICSCAFAL 646
DB 588 IKDALYFLDDTAKMVPVMTEDINYAISAKTLKL 620
RESULT 5
VLCS_MOUSE
ID VLCS_MOUSE STANDARD; PRT; 620 AA.

AC 035488; 070550;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE VERY-LONG-CHAIN ACYL-COA SYNTHETASE (EC 6.2.1.-) (VERY-LONG-CHAIN-
DE FATTY-ACYL-COA LIGASE)
GN FACVLI OR VLACS OR VLCS OR SLC27A2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Liver;
RA Kemp S., Lu J.-F., Smith K.D.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BA1B/C; TISSUE=Liver;
RX MEDLINE=98218572; PubMed=9559670;
RA Berger J., Truppe C., Neumann H., Forss-Petter S.;
RT "cDNA cloning and mRNA distribution of a mouse very long-chain acyl-
RT CoA synthetase";
RL FEBS Lett 425:305-309(1998).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PEROXISOMAL.
CC -!- TISSUE SPECIFICITY: STRONG EXPRESSION IN LIVER AND KIDNEY, LOW
CC EXPRESSION IN BRAIN AND TESTIS, NO EXPRESSION IN SKELETAL MUSCLE
CC AND SPLEEN.
CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

EMBL; AF033031; AAB87982.1; -.
DB EMBL; AJ223958; CAAL1687.1; -.
DR MGD; MGI:1347099; SLC27a2.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
KW Ligase; Fatty acid metabolism; Peroxisome; Transmembrane.
FT TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 107 127 POTENTIAL.
FT TRANSMEM 262 282 POTENTIAL.
FT CONFLICT 35 35 Q -> R (IN REF. 2).
SQ SEQUENCE 620 AA; 70366 MW; 77C98BD0DE3B9FFB CRC64;

Query Match 30.6%; Score 1032.5; DB 1; Length 620;
Best Local Similarity 38.7%; Pred. No. 6.1e-73;
Matches 241; Conservative 97; Mismatches 254; Indels 31; Gaps 12;

QY 34 VYVSGGWRFLRIV-----CKTARRDL---FGLSVLIRVLELRRHORAGHTIPRIFOAVV 86
DB 5 IYTVLAGLLFLPLNLNCCPFYFDIGYFLKAAVGRVRSYQGRRTILRAFLPKAR 63
QY 87 ROPERLALVDAGTGCWTFQAOLDAYSNAVAN-LFRQLGFAPGDVVAIFLEGRPEFVGLW 145
DB 64 RKTTPKFLU--FRDETLTYAQRDRSNQVARALHDHGLRGDCVALLMGNEPAYVWLW 121
QY 146 GLAKAGMAEALLNVLNREPLAFCLGTSKAKALIPFGGEMVAEVSGLKSLTK--F 203
DB 122 LGLKLGCMACLNINRAKSLHLCPCCGAKVLLASPDQAEAEALPTLKKDVAVSFY 181
QY 204 CSGDLGPEGILPDLTHLLDPLLEKASTAPLAQIPSKGMDORLF-----YIYTSCTTGLPKA 258
DB 182 VSRSTNTNGV-----DTILDRKVDGVSAPETPESWRSEVTFPTTPAVIYTSCTTGLPKA 234

QY 259 AIVVHSRYRMAAFHHAYRMOADVLYDCLPLYHSAGNLIGVQCGLIYGLTVLVRKFS 318
 Db 235 ATINHHRLRYGTGLAMSS-CITADQVLYITMPLYSAAALMIGLHGCIVVGATIALRSKFS 293
 QY 319 ASRFWDCCIKYNTVVOYIGEICRYLLKOPVREARHRVRLAVGNGLRPAIWEFEFRF 378
 Db 294 ASQFWDCCRYNTVVIQYIGELLRYLCNTPKPNDROHVKKALGNLGRDGVWREIFRF 353
 QY 379 GVRIQGEFGATCNCNIANDCKVSCGFNSRILPHVPIRLVKVNETMELLRDAQGL 438
 Db 354 GDHIVFEYASTGNIGFVNPRIKIGAVGRANLYRKVARELYIKYDVDEKDFVRDANGY 413
 QY 439 CIPQAGEPGLLYGOINQODPLRRFDGYV--SESATSKIAHSVFESKSDSAYLSGDVLYMD 497
 Db 414 CIKVPKGEVGLLVCKITQLTTP---FIGYAGGKTQTEKKLRLDVFKKGDIFYNSGDLIMID 470
 QY 498 ELGYMYFRDRSGDTFRWRGENVSTTEVEGVLRLGQTDVAVYGVAVPVEGKAGMAVA 557
 Db 471 RENFVYFHDVGRVDTFRMKGENVATTEVADIVGLDVEEVNNGVVPVPGHEGRIGMASLK 530
 QY 558 DPHSL-LDPNIAIYQELQKVLAPYARPIFLRLLPQVDTTCTGTEKIQTRLOREGDFDRQTS 616
 Db 531 IKENYFENGKLPQHTAEVLPYARPRFLRIQDTIHTGFRKRVTLMEEGFNPTVIKD 590
 QY 617 RLFFLDLKOGHYLPPLNEAVYTRI 639
 Db 591 TLYFMDDAEKTFFVPTENIYNAI 613

RESULT 6

FAT1_YEAST
 ID FAT1_YEAST STANDARD; PRT; 623 AA.
 AC F38225;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 20-OCT-1994 (Rel. 30, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROBABLE LONG-CHAIN FATTY ACID TRANSPORT PROTEIN.
 GN FAT1 OR YBR041W OR YBR041L.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Andre B., Cziepluch C., Hein C., Jauniaux J.C., Urrestarazu A.,
 RA Viissers S.;
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RP CHARACTERIZATION.
 RC STRAIN=W303A;
 RC MEDLINE=97236810; PubMed=9079682;
 RX Faergeman N.J., Dirusso C.C., Elberger A., Knudsen J., Black P.N.;
 RA "Disruption of the Saccharomyces cerevisiae homologue to the murine
 RT fatty acid transport protein impairs uptake and growth on long-chain
 RT fatty acids.";
 RL J. Biol. Chem. 272:8531-8538(1997).
 CC -!- FUNCTION: MAY BE INVOLVED IN LONG-CHAIN FATTY ACIDS UPTAKE, AND
 CC THUS MAY PLAY A PIVOTAL ROLE IN REGULATING THEIR ACCESSIBILITY
 CC PRIOR TO METABOLIC UTILIZATION. MAY PLAY AN IMPORTANT ROLE IN
 CC UPTAKE OF THESE HYDROPHOBIC COMPOUNDS UNDER CONDITIONS WHERE FATTY
 CC ACID SYNTHESIS IS COMPROMISED.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or
 or send an email to license@isb-sib.ch).

CC

DR EMBL; Z35910; CAA84983.1; -.
 DR PIR; S45899; S45899.
 DR SGD; S0000245; RAT1.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00501; AMP-binding; 1.
 DR PROSITE; PS00455; AMP_BINDING; 1.
 KW Lipid transport; Transmembrane.
 FT TRANSMEM 6 26 POTENTIAL.
 FT TRANSME 54 71 POTENTIAL.
 FT TRANSMEM 149 169 POTENTIAL.
 FT TRANSMEM 293 313 POTENTIAL.
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 534 534 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 623 AA; 71697 MW; 0AE02700B60C8CFE CRC64;

Query Match

Best Local Similarity 20.7%; Score 699; DB 1; Length 623;
 Matches 190; Conservative 87; Mismatches 234; Indels 56; Gaps 19;

QY 66 LELRRHORAGHTIPRIQAVVQROPERLA-----LYDAGTGECWTEAQLDAYSNAYNL 119
 Db 69 IDVRHRFQNW---YLFIKVQOQNGDHLAISYTRPMAEKGEFQLETFTYETY-NIVLR 124
 QY 120 FRQLGF----APGDVVAIFLEGREPFVGLWGLAKAGMAEALNVNLRREPLAFCLGTSG 175
 Db 125 SHILHFDYVQAGDYVAIDCTNKPFLFVFLWLSLWNIGAIPAFLNNTKGPPLVHSLKSN 184
 QY 176 AKALIFGEMVAOVAEVSGLHKLKSLKFCSDGLGPE---GILPDTLLDPLLEKASTAPL 232
 Db 185 ITQVFDIPDASNP IRESEEEIKNAL-----PDYKLNYLEEQDLMHLLASQSPEFL 235
 QY 233 AQIPSK---GMDD--RLFYIVTSGTGLPKAAIVVHSRYRMAAFG---HHAYRQOAA 283
 Db 236 QQDNVTRPLGLTDFKPSMLIYTSCTGLPKSATM---SWRKSQVQGVGHVLMHTNES 291
 QY 284 VLYDCLPLYISAGNIIGVQCGLYGLVLRKFSARFWDCCIKYNTVVOYIGEICRY 343
 Db 292 TVFTAMPFLHSTALLGACAILSHGGCALSLSHRKFSASTFWKVLGTGATHIQVGVVCRY 351
 QY 344 LLKOPVREARRHRVRLAVGNGLRPAIWEFEFRFVRQIGEFYGATE---CNCSTANMD 400
 Db 352 LLHTPISKYEMHKVKVAYGNGLRPDIWQDFRRFRNIEVIEGYAATEAFATTTQKGD 411
 QY 401 GKVSCGFNSRIPLPHVYPIR--LVKV--NEDTMELLRDAQGLCIPQAGBPG--LLVGQIN 455
 Db 412 FGIGACRNYGTIIQWFLSFOOTLVRMDPNDSV-IYRNSKGFCEVAPVGPGEMLKRIFF 470
 QY 456 QQDPLRRFDGYVSES-ATSKIAHSVFESKSDSAYLSGDVLMDELGYMYFRDRSGDTFRW 514
 Db 471 PKPETSFOGLYGNAKETKSKVVRDVRFRGDWYRCGDLKKADEYGLIWFYLDKMGDTFRW 530
 QY 515 RGENVSTTEVEGVL--SRLLGQTDVAVYGVAVPVEGKAGMAVADPHSLSDPNA---I 568
 Db 531 KSENVSTTEVEDQLTASNKEQYAOVLVVGKVPKYEGRAGFAVIKLTDSLDTAKTKLL 590
 QY 569 YQELQKVLAP-YARPIFLRLLPQVDTT 594
 Db 591 NDSLSRLNLPYAMPLFVKFVDEIKWT 617

RESULT 7

CAIC_ECOLI
 ID CAIC_ECOLI STANDARD; PRT; 522 AA.
 AC P31552;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROBABLE CROTONOBTAININE/CARNITINE-COA LIGASE (EC 6.3.2.-).
 GN CAIC OR B0037.

OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=O44 K74;
RX MEDLINE=95115548; PubMed=7815937;
RA Eichler K., Bourgis F., Buchet A., Kleber H.-P.,
RA Mandrand-Berthelot M.-A.,
RT "Molecular characterization of the cai operon necessary for carnitine
metabolism in Escherichia coli.",
RL Mol. Microbiol. 13:775-786(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92334977; PubMed=1630901;
RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
RA Isono K., Mizobuchi K., Nakata A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
the 0-2.4 min region.",
RL Nucleic Acids Res. 20:3305-3308(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1453-1474(1997).
CC -!- FUNCTION: COULD CATALYZES THE TRANSFER OF COA TO CROTONOBETAINE
OR CARNITINE.
CC -!- PATHWAY: CARNITINE METABOLISM.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
(POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X73904; BAA52113.1; -.
DR EMBL; D10483; BAA01313.1; -.
DR EMBL; AE000114; AAC73148.1; -.
DR PIR; S40558; S40558.
DR HSP; P08659; ILCI.
DR EcoGene; EGI11558; caic.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS00455; AMP_BINDING; 1.
KW Ligase; Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 77 97
FT POTENTIAL.
FT TRANSMEM 231 251
FT POTENTIAL.
FT VARIANT 103 103
FT C -> R (IN STRAIN O44 K74).
FT Q -> R (IN STRAIN O44 K74).
FT I -> S (IN STRAIN O44 K74).
FT V -> A (IN STRAIN O44 K74).
FT I -> V (IN STRAIN O44 K74).
FT V -> K (IN STRAIN O44 K74).
FT A -> V (IN STRAIN O44 K74).
FT R -> C (IN STRAIN O44 K74).
FT D -> G (IN STRAIN O44 K74).
FT A -> T (IN STRAIN O44 K74).
FT L -> LDL (IN STRAIN O44 K74).

SQ SEQUENCE 522 AA; 59089 MW; LC54D4003B130671 CRC64;
Query Match 11.6%; Score 390.5; DB 1; Length 522;
Best Local Similarity 25.7%; Pred. No. 6.7e-23;
Matches 138; Conservative 84; Mismatches 251; Indels 53; Gaps 16;
QY 92 RLALV---DAGTGECWTFQAQDAYSNANVAFQGFAGFDVVAIFLEGRPEFVGLWGL 148
Db K T A L I C E S S G G V N R Y S Y L E L N Q E I N R T A N L E V T L I R K G D K V A L H L D N C P E F I C W F G L 87
QY 149 AKAGMEAAALNNVLRREPLAFCLGTSGAKALFEGGEMAAVAEVSCHGKSLIKKCSGDL 208
Db K T A L I C E S S G G V N R Y S Y L E L N Q E I N R T A N L E V T L I R K G D K V A L H L D N C P E F I C W F G L 87
QY 88 AKIGATWVPINARLCEESAWILQNSOACLLVTSAQFYPMYQOIQOEDATQRHICLDIV 147
Db K T A L I C E S S G G V N R Y S Y L E L N Q E I N R T A N L E V T L I R K G D K V A L H L D N C P E F I C W F G L 87
QY 209 ---GPEGILPDTHLPLLEKASTAPLAQIPSKGMDRLFIYITSGTTPKAAAVVHSR 265
Db K T A L I C E S S G G V N R Y S Y L E L N Q E I N R T A N L E V T L I R K G D K V A L H L D N C P E F I C W F G L 87
QY 266 YRMAAFGHHH---YRMAADVLYDCPLIYHSAGNIIGVQCCLLYGLTVLVRKKFSASRF 322
Db K T A L I C E S S G G V N R Y S Y L E L N Q E I N R T A N L E V T L I R K G D K V A L H L D N C P E F I C W F G L 87
QY 323 WDDCIKYNCTVVOYIGEICRYLLKQPVREARRHVR-LAVGNGLRPAIWEETREGRV 381
Db K T A L I C E S S G G V N R Y S Y L E L N Q E I N R T A N L E V T L I R K G D K V A L H L D N C P E F I C W F G L 87
QY 382 QIGEFYGATGECNCSIAN-----MDGKVGSCFCNRSIL-PHYVPIRLVKVNEDTME 430
Db K T A L I C E S S G G V N R Y S Y L E L N Q E I N R T A N L E V T L I R K G D K V A L H L D N C P E F I C W F G L 87
QY 431 LLRDAOGLCIPCAOGEPLLGVQINOQDPLRRFDGVVSESATSKIAHSVFKSGDSAYLS 490
Db K T A L I C E S S G G V N R Y S Y L E L N Q E I N R T A N L E V T L I R K G D K V A L H L D N C P E F I C W F G L 87
QY 491 GDVLMDELGYMFRDRSGDTFRWGENVSTTEVEGLSKLLGQTDVAVYGVAVPGVEGK 550
Db K T A L I C E S S G G V N R Y S Y L E L N Q E I N R T A N L E V T L I R K G D K V A L H L D N C P E F I C W F G L 87
QY 551 AGMA-AVADPHSLDPAIYQELQKVLAPYARPIFLRLPQVDVTGTFKIQTRIQ 605
Db K T A L I C E S S G G V N R Y S Y L E L N Q E I N R T A N L E V T L I R K G D K V A L H L D N C P E F I C W F G L 87
RESULT 8
PPS2_BACSU STANDARD; PRT; 2560 AA.
AC P39846;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PEPTIDE SYNTHETASE 2.
GN PPSB OR PPS2.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=95227362; PubMed=7711903;
RA Tognoni A., Franchi E., Magistrelli C., Colombo E., Cosmina P.,
RA Grandi G.;
RT "A putative new peptide synthase operon in Bacillus subtilis: partial
characterization.",
RL Microbiology 141:645-648(1995).
CC -!- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTHETINE
(POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; 234883; CAA84361.1; -
 EMBL; 299113; CABI3716.1; -
 HSSP; P14687; IAMU.
 Subtilist; BG10971; ppsB.
 InterPro; IPR000873; AMP-bind.
 InterPro; IPR001242; DUF4.
 InterPro; IPR003880; Phosphopant_attach.
 Pfam; PF00301; AMP-binding; 2.
 Pfam; PF00668; Condensation; 3.
 Pfam; PF00550; pp-binding; 2.
 PRINTS; PR00154; AMPBINDING.
 PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
 PROSITE; PS00455; AMP-BINDING; 2.
 PROSITE; PS00075; ACP_DOMAIN; 2.
 Complete proteome.
 Multifunctional enzyme; Ligase; Repeat; Phosphopantetheine;
 KW
 DOMAIN 970 1037 ACYL CARRIER (ACP) 1.
 FT BINDING 2007 2077 ACYL CARRIER (ACP) 2.
 FT BINDING 2041 2041 PHOSPHOPANTHETHEINE (POTENTIAL).
 SQ SEQUENCE 2560 AA; 290161 MW; 2DD2442D11B6E942 CRC64;

Query Match 8.9%; Score 301.5; DB 1; Length 2560;
 Best Local Similarity 23.7%; Pred. No. 5.2e-15;
 Matches 130; Conservative 82; Mismatches 251; Indels 85; Gaps 13;

QY 77 TIPRIQAVVQPERLALVDAGTGECWTFADLADAYNAVANLRLQGLFAPGDVVAIFLE 136
 DB 470 TIPQLFEQAHKTPRAALKMG-NECWYRQLQVRAQIAHALIEKVGSGDIVVMWG 527
 QY 137 GRPEFVGLWGLAKAGMEALLNVNLRREPLAFCLGTSGAKALIFGSGEMVAABVSGHL 196
 DB 528 RSMEMPAALLGIKAGKAGMAYPLDPHFAERLSFLKDKSOAQLLEEDLIS- 578
 QY 197 GSKLIKFCSGDLGPGSILLPDTHLLDPLKEASTAPLAQIPSKGMDRLFIYTGTTGLP 256
 DB 579 -----LIPPSEGTITTEHTESYQTEAPNPPGDLAYLIYTGTTGRP 622
 QY 257 KAAIVVHSRYRMAAFHHAYRMAADVLDCPLYSAGNTIGVGCLLYGLVLRKK 316
 DB 623 KGVLDHHCIAANTLQWRREYSMTQDISLH-LFSYVFDGCVTSLFTPLLSGACVLLTD 681
 QY 317 FSASRFWDCKINYCTVVOY-----IGEICRYLLKQPVREARRHRVRLAVGNGLRPA 369
 DB 682 DEAK-----DVLAKRKEIARYKVSHMIIVPSLYRVLLEVMTADAKSLRIVTFAGEAVTPD 737
 QY 370 IWEETERFRGVQICEFYGATECNCSIANMDGKVGSCGFNSRILPHV-----YPIR 420
 DB 738 LLELQNIQICPSANLANEGPTE--NSVA-----PTILRLHKKERITIGPHIR 783
 QY 421 LVKV---NEDTMEILLRDAQGLICPCQAGEPGLLVGQINQDPLRRFRDGYVSESATKIA 477
 DB 784 NTKVPVLHGNQMOPIGAAGELCI-----SGAGLARGYKQOE-----LTQAFS 827
 QY 478 HVSFSGDSAYLSDGVLDVMDLGYWYFRDRSGDTFRNGENVSTVEGVLRLGGLQTDV 537
 DB 828 DHPFLEGGRLRYTGDAGRFLPDGTGTYIGRFDQDVQKIRGYRIELREIETVLRAPGVKEA 887
 QY 538 AVYGVAVPGVKGAGMAAVADP-HSLDDPNNAIYQELQKVLAPYARPIFLRLPQVDITGT 596
 DB 888 AVLARDVSAEKELVAYLVPEKGNLSDP---LYOHLAGTLPSYMPIASIINISOMPLTSS 944
 QY 597 FKIQKTRL 604
 DB 945 GKLDRFAL 952

RESULT 9

4CLL_PETCR
 ID 4CLL_PETCR STANDARD; PRT; 544 AA.
 AC P14912;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE 4-COUMARATE--COA LIGASE 1 (EC 6.2.1.12) (4CL 1) (4-COUMAROYL-COA
 DE SYNTHASE 1).
 GN 4CL1 OR 4CL-1.
 OS Petroselinum crispum (Parsley) (Petroselinum hortense).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
 RN NCBI_TaxID=4043;
 RX [1]
 SEQUENCE FROM N.A.
 RX MEDLINE=9005119; PubMed=3169018;
 RA Lozoya E., Hoffmann H., Douglas C., Schulz W., Scheel D.,
 RA Hahlbrock K.;
 RA "Primary structures and catalytic properties of isoenzymes encoded by
 RT the two 4-coumarate:CoA ligase genes in parsley.";
 RL Eur. J. Biochem. 176:661-667(1988).
 RN [2]
 SEQUENCE OF 1-8 FROM N.A.
 RA Douglas C., Hoffmann H., Schulz W., Hahlbrock K.;
 RT "Structure and elicitor or U.V.-light-stimulated expression of two
 RT 4-coumarate:CoA ligase genes in parsley.";
 RL EMBO J. 6:1189-1195(1987).
 CC -|- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA -> AMP + DIPHOSPHATE +
 CC 4-COUMAROYL-COA.
 CC -|- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -|- INDUCTION: TRANSCRIPTIONALLY ACTIVATED BY BOTH FUNGAL ELICITOR AND
 CC ULTRAVIOLET IRRADIATION.
 CC -|- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

Query Match 8.9%; Score 298.5; DB 1; Length 544;
 Best Local Similarity 23.9%; Pred. No. 1.1e-15;
 Matches 128; Conservative 95; Mismatches 231; Indels 81; Gaps 18;

QY 48 CKTARDRLGSLVLR--IRVLELRHRRHAGHTIPRFQAVVQPERLALVDAGTGCWT 105
 DB 4 CVAPKEDLLFRSLDPDIYIPKHLPLHYTCFENISKY-----GDKSLINGATGETPT 55
 QY 106 FAQLDAYSNANLRLQGLFAPGDVVAIFEGRPVGLWGLAKAGMEALLNVNLRRE 165
 DB 56 YSQVELLSRKVASGLNKLGIQOGDTIMLLPNSPEYFFAFGLGASVGAISTWANFFISA 115
 QY 166 PLAFCLGTSGAKALIFGSGEMVAABVSGHLGSKSLIKFCSGDLGPGGLIPDTHLLDPLK 225
 DB 116 EVIKQLKASQAKLII---TQACYVDKVKDYAAAEKNIQIICIDDAPOQDCLHFSKLME---A 169

QY 226 EASTAPLAQIPKSGMDRLFIYITSGTGLPKAAIVVH-----SRYRMAAFGHAYRMO 280
 Db 170 DESEMPVINS---DDVVALLPYSSGTTGLPKGVLTHKGLVTSVAQVQDGNPNLY-MH 225
 QY 281 ADVLYDCLPLYH--SAGNIIGVQCGLIYGLTVLVRKFSASRWDCCIKYKNCIVVQYIG 338
 Db 226 SEDVMICILPLPHIYSLNVLCCG--LRAGVTILIMQKFDIVPFLQLQYKVTIGPFPV 283
 QY 339 EICRYLLKOPVREARRHVRLLA-----VGNGLRPAIWEETFERFGVQIGFYGATEC 392
 Db 284 PIVLAIKSPVVDKYLSSVRTVMSGAAPLKELEDVAKFPN---AKLGQGYGTEA 339
 QY 393 NCSIA-----NMDGKVGSCGFNSRILPHYPYIRLVKVNEDTMELLRDAGGLCIP-CQ 443
 Db 340 GPVLAMCLAFAPKEPIKSGAGC-----TWVRNAEMKIVDPETNASLPNRQ 385
 QY 444 AGEPLLVGOINQODPLRRFDGYVSESATSKIAHSVSKGDSAYL-SGDVLVMDLGYM 502
 Db 386 RGEICIRGDOI-----MKGYLNDPESTRT-----TIDEGWLHTGDIIDDDEL 431
 QY 503 YFDRSGDTFRWRGENVSTTEVEGVLRLGQTDVAVYGVAVPGVGEKAGMAAVA 557
 Db 432 FIVDRLEIKIYKGFQVAPAELEALLTHPTISDAAV---VPMIDERAGEVPVA 482

RESULT 10

4CL2_PETCR STANDARD; PRT; 544 AA.
 ID 4CL2_PETCR
 AC P14913;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE 4-COUMARATE-COA LIGASE 1 (EC 6.2.1.12) (4CL 1) (4-COUMAROYL-COA
 DE SYNTHASE 1).
 GN 4CL2 OR 4CL-2.
 OS Petroselinum crispum (Parsley) (Petroselinum hortense).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asterales; euasterids II; Apiales; Apiaceae; Petroselinum.
 OX NCBI_TaxID=4043;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89005119; PubMed=3169018;
 RA Lozoya E., Hoffmann H., Douglas C., Schulz W., Scheel D.,
 RA Hahlbrock K.;
 RT Primary structures and catalytic properties of isoenzymes encoded by
 RT the two 4-coumarate:CoA ligase genes in parsley.";
 RL Eur. J. Biochem. 176:661-667(1988).
 RN [2]
 RP SEQUENCE OF 1-8 FROM N.A.
 RA Douglas C., Hoffmann H., Schulz W., Hahlbrock K.;
 RT "Structure and elicitor or U.V.-light-stimulated expression of two
 RT 4-coumarate:CoA ligase genes in parsley.";
 RL EMBO J. 6:1189-1195(1987).
 CC -!- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
 CC 4-COUMAROYL-COA.
 CC -!- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -!- INDUCTION: TRANSCRIPTIONALLY ACTIVATED BY BOTH FUNGAL ELICITOR AND
 CC ULTRAVIOLET IRRADIATION.
 CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X13325; CAA31697.1; -;
 CC EMBL; X05351; CAA28960.1; -;

DR PIR; S01667; S01667.
 DR PIR; S15695; S15695.
 DR HSP; P08659; ILCI.
 DR InterPro: IPR000873; AMP-binding.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP-BINDING; 1.
 KW Ligase; Phenylpropanoid metabolism; Multigene family.
 SQ SEQUENCE 544 AA; 59783 MW; B477965C68F8C534 CRC64;

Query Match 8.8%; Score 297.5; DB 1; Length 544;
 Best Local Similarity 23.9%; Pred. No. 1.3e-15;
 Matches 128; Conservative 95; Mismatches 231; Indels 81; Gaps 18;

QY 48 CKTARRDLFGLSVL--IRVRLRLRRHQRAGHTTIPRIFOAVVORQPERLALVDAGTGECWT 105
 Db 4 CVAPKEDLIFRSKLPDIYIPKHLPLHYCFENISKV-----GDKSLINGATGETPT 55
 QY 106 FAQDDAYSNAVANLFRQLGAPQGVVAIFLEGPRFVGLWGLAKAGMAALLNVNLRRE 165
 Db 56 YSQVELLSRKVASGLNKLGIQGGDTIMLLPNPSPEYFFAFLGASYRGALSTMANPFTSA 115
 QY 166 PLAFCLGTSGAKALIFGEMVAAVAESVGHGKSLIKFCGDLGPEGILPDTHLLDPLLK 225
 Db 116 EVIKQLKASLAKLII---TQACVYDKVKDYAAEKNTIICIDDAPODCLHFSKLME---A 169
 QY 226 EASTAPLAQIPKSGMDRLFIYITSGTGLPKAAIVVH-----SRYRMAAFGHAYRMO 280
 Db 170 DESEMPVIDS---DDVVALLPYSSGTTGLPKGVLTHKGLVTSVAQVQDGNPNLY-MH 225
 QY 281 ADVLYDCLPLYH--SAGNIIGVQCGLIYGLTVLVRKFSASRWDCCIKYKNCIVVQYIG 338
 Db 226 SEDVMICILPLPHIYSLNVLCCG--LRAGVTILIMQKFDIVPFLQLQYKVTIGPFPV 283
 QY 339 EICRYLLKOPVREARRHVRLLA-----VGNGLRPAIWEETFERFGVQIGFYGATEC 392
 Db 284 PIVLAIKSPVVDKYLSSVRTVMSGAAPLKELEDVAKFPN---AKLGQGYGTEA 339
 QY 393 NCSIA-----NMDGKVGSCGFNSRILPHYPYIRLVKVNEDTMELLRDAGGLCIP-CQ 443
 Db 340 GPVLAMCLAFAPKEPIKSGAGC-----TWVRNAEMKIVDPETNASLPNRQ 385
 QY 444 AGEPLLVGOINQODPLRRFDGYVSESATSKIAHSVSKGDSAYL-SGDVLVMDLGYM 502
 Db 386 RGEICIRGDOI-----MKGYLNDPESTRT-----TIDEGWLHTGDIIDDDEL 431
 QY 503 YFDRSGDTFRWRGENVSTTEVEGVLRLGQTDVAVYGVAVPGVGEKAGMAAVA 557
 Db 432 FIVDRLEIKIYKGFQVAPAELEALLTHPTISDAAV---VPMIDERAGEVPVA 482

RESULT 11

SRFL_BACSU STANDARD; PRT; 3587 AA.
 ID SRFL_BACSU
 AC P27206;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SURFACTIN SYNTHETASE SUBUNIT 1.
 GN SRFAA OR SRFAI OR SRFA.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=93181186; PubMed=8441623;
 RA Fuma S., Fujishima Y., Corbell N., D'Souza C., Nakano M.,
 RA Zuber P., Yamane K.;
 RT "Nucleotide sequence of 5' portion of srfa that contains the region
 RT required for competence establishment in Bacillus subtilis.";
 RL Nucleic Acids Res. 21:93-97(1993).

Db 2604 PSEMSARCLGVLKAGAAFPIDPDYDQRIEYLQDSGAKL----- 2646
 QY 197 GSKLIKFCGDLGPBGI-LPDTLLDPLKKEASTAPLAQIPSKGMDR-----L 244
 Db 2647 -----LKQGISVPSYSGDVILLDSKRTILS-LPLDDEENPETAVTAENLA 2694
 QY 245 FYIYTSGTGTPKAAIVVHSRYRMAAFHHAYRMAQAD-----VLYDCUPLY 292
 Db 2695 YMIYTSGTGTPKGVVHEHALVNLFCFHHDAFSNTAEDRSKYAGFGFASIWEMEPPT 2754
 QY 293 HSAGNIIGVQCGLIYGLVLRKFKFSASFWDCCIKNCTVQYIGEICRVLKQPVREA 352
 Db 2755 TIGAEHLVIEAIRLDIVRL-----NDYFETNGVTITFL-----PTQLA 2793
 QY 353 ER-----RHRVLAVNGRLPALIEBETFRFGVQRIEGEYATECNCSTANMDGKVG 405
 Db 2794 EQFMELENTSLRVLLTGDKLRAVKKPYT-----LVNNYGPTE-NTVVAT----- 2838
 QY 406 CGFNSRILPHVYPIRLVKNEDTMELLRDAQGLCIP-----COAGEPGLLVGQINQD 458
 Db 2839 ---SAEIHPEGSLSIGRAIANTRYILGEGNQVQPEGVAGELCVAGR-GLARGYLNRED 2894
 QY 459 PLRRDGGVYESATSKTAHVSFKGDSAYLSGDVLYMDELGYMYFRDRSGDTFRWRGEN 518
 Db 2895 -----ESPKRVADPFVPGERMYRTGDLVKTWGGIEYI-GRIDQVKVGRYR 2941
 QY 519 VSTTEVEGVLRLGQTDVAVYGVAVPGVEGKAGMAAADVPHSLDPAIYQELQKVLAP 578
 Db 2942 IELSEIYQALQSEVQDAAVTAVKDKG--GNTAIAAVYTPES-ADTEALKSALKETLPD 2998
 QY 579 YARPIFLRLQVDTTGTFTKIQKTRL 604
 Db 2999 YMPAFVWTLNLPVTANGKVDKRAL 3024

RESULT 12

4CL1_TOBAC STANDARD; PRT; 547 AA.
 AC 024145;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE 4-COUMARATE--COA LIGASE 1 (EC 6.2.1.12) (4CL 1) (4-COUMAROYL-COA
 DE SYNTHASE 1).
 GN 4CL1.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96416441; PubMed=8819324;
 RA Lee D., Douglas C.J.;
 RT "Two divergent members of a tobacco 4-coumarate:coenzyme A ligase
 (4CL) gene family. cDNA structure, gene inheritance and expression,
 and properties of recombinant proteins.";
 RT Plant Physiol. 112:193-205(1996).
 CC -!- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA -> AMP + DIPHOSPHATE +
 4-COUMAROYL-COA.
 CC -!- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U50845; AAB18637.1; -
 DR HSP; P08659; 1LC1.
 DR Mengel; 24477; Nicta;1179;24477.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00501; AMP-binding; 1.
 DR PROSITE; PS00455; AMP_BINDING; 1.
 KW Ligase; Phenylpropanoid metabolism; Multigene family.
 SQ SEQUENCE 547 AA; 59842 MW; 154DF6D684E3F51B CRC64;
 Query Match 8.6%; Score 290.5; DB 1; Length 547;
 Best Local Similarity 24.7%; Pred. No. 4.6e-15;
 Matches 121; Conservative 82; Mismatches 215; Indels 71; Gaps 17;
 QY 92 RLALVDAGTCEWTFCAQLDAYSNAVANLFRQLGFGPDGVVAIFLEGRPEFVGLWGLAKA 151
 Db 46 RPLCLINGANDQIVTYAEVELTCRKVAVGLNKLGIOQKDFIMILLPNSPEFVFAFGASVL 105
 QY 152 GMPEALLNVLNREPLAFCLGTSGAKALIFGGEWAANAESVHGLKSLKFCSDGLGPE 211
 Db 106 GAISTMANPLFTPAEYVVKQAKASAKIIIT---TQSCFVGKVDYASENDVKVICIDSAPE 162
 QY 212 GILPDTHLLDPLKKEASTAPLAQIPSKGMDRLFYITSGTGLPKAAIVHHSRYRMAA 271
 Db 163 GCL---HFSELQTSDHEIPEVKIQP---DDVVALPYSSGTTGLPKGYMLTHKGLVTSVA 216
 QY 272 ---FGHHA-YRMOAADVLYDCLPLXH--SAGNIIGVQCGLIYGLVLRKFKFSASRFDWD 325
 Db 217 QQVDGENANLYHSEDLVLCVLPFLPHIYSLNLLCG--LRVGAAILIMQKFDIAPFLEL 274
 QY 326 CIKYNCTVQYIGEICRYLLKQPVREAERHVRLA-----VGNGLRPAIWEETEREG 379
 Db 275 IQKYKVSIGTFVPIVLAITAKSPIDVSDSSVTWMSGAAPLGKLEDAVTKFPN--- 331
 QY 380 VROIIEFYGATECNCSTIA-----NMDKVGSCGFNSRILPHVYPIRLVKNEDTMEL 431
 Db 332 -AKLGQGYGTEAGPVLAMCLAFAPKPEFDIKSGAG-----TVVRNAEMKI 376
 QY 432 LRDAQGLCIP-COAGEPGLLVGQINQDPLRRPDGVSE-SATSKIAHSVFSKGSAYL 489
 Db 377 VDPDTGCSLPRNQPEICIRGQI-----MKGYLNDPEATTRTI-----DKEGWL 421
 QY 490 -SGDLVLMDELGYMYFRDRSGDTFRWRGENSVTEVEGVLSRLGQTDVAVYGVAVPGVE 548
 Db 422 HTGDIGFIDEDELFIYDRKLKELIKYGFQVAPAEIEALLNHPNISDAV-----VPMKD 477
 QY 549 GKAGMAAVA 557
 Db 478 EQAGEVPVA 486
 RESULT 13
 SRF2_BACSU
 ID SRF2_BACSU STANDARD; PRT; 3587 AA.
 AC 004747;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SURFACTIN SYNTHETASE SUBUNIT 2.
 GN SRFAB OR SRF2 OR COML.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE OF 1-3077 FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=93181186; PubMed=8441623;
 RA Fuma S., Fujishima Y., Corbell N., D'Souza C., Nakano M.M.,
 RA Zuber P., Ramane K.;
 RT "Nucleotide sequence of 5' portion of srfA that contains the region
 RT required for competence establishment in Bacillus subtilis.";

RL Nucleic Acids Res. 21:93-97(1993).
 RP [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=168/JH642;
 RX MEDLINE=93360813; PubMed=8355609;
 RA Cosmina P., Rodriguez F., de Ferra F., Grandi G., Perego M.,
 RT "Sequence and analysis of the genetic locus responsible for surfactin
 synthesis in *Bacillus subtilis*.";
 RL Mol. Microbiol. 8:821-831(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97124189; PubMed=8969502;
 RA Yamane K., Kumano M., Kurita K.;
 RT "The 25 degrees-36 degrees region of the *Bacillus subtilis*
 chromosome: determination of the sequence of a 146 kb segment and
 identification of 113 genes.";
 RL Microbiology 142:3047-3056(1996).
 RN [4]
 RP SEQUENCE OF 514-800 FROM N.A.
 RC STRAIN=ATCC 21332;
 RX MEDLINE=92290255; PubMed=1601288;
 RA Borchert S., Fatli S.S., Marahiel M.A.;
 RT "Identification of putative multifunctional peptide synthetase genes
 using highly conserved oligonucleotide sequences derived from known
 synthetases.";
 RL FEMS Microbiol. Lett. 71:175-180(1992).
 CC -!- FUNCTION: THIS PROTEIN IS A MULTIFUNCTIONAL ENZYME ABLE TO
 ACTIVATE AND POLYMERIZE THE AMINO ACIDS LEU, GLU, ASP AND VAL.
 CC ACTIVATION SITES FOR THESE AA CONSIST OF INDIVIDUAL DOMAINS.
 CC -!- COFACTOR: CONTAINS 3 COVALENTLY BOUND PHOSPHOPANTHETHEINES.
 CC -!- PATHWAY: NON-RIBOSOMAL BIOSYNTHESIS OF THE CYCLIC PEPTIDE
 ANTIBIOTIC SURFACTIN.
 CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 FAMILY
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D13262; BAA02523.1; -;
 DR EMBL; X70356; CAA49817.1; -;
 DR EMBL; D50453; BAA08983.1; -;
 DR EMBL; X65835; CAA46678.1; -;
 DR EMBL; Z99105; CAB12143.1; -;
 DR HSSP; P14687; IAMU.
 DR Subtilisin; BG10169; srfAB.
 DR InterPro; IPR000873; AMP-bind.
 DR InterPro; IPR001242; DUF4.
 DR InterPro; IPR003680; Phosphopant_attach.
 DR Pfam; PF00501; AMP-binding; 3.
 DR Pfam; PF00668; Condensation; 4.
 DR Pfam; PF00550; pp-binding; 3.
 DR PRINTS; PR00154; AMPBINDING.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 3.
 DR PROSITE; PS00455; AMP BINDING; 3.
 DR PROSITE; PS00075; ACP_DOMAIN; 3.
 KW Ligase; Antibiotic biosynthesis; Phosphopantetheine; Sporulation;
 KW Multifunctional enzyme; Repeat; Complete proteome.
 FT REPEAT ? 1040 DOMAIN 1 (VAL-ACTIVATING).
 FT REPEAT ? 2096 DOMAIN 2 (ASP-ACTIVATING).
 FT REPEAT ? 3114 DOMAIN 3 (D-LEU-ACTIVATING).
 FT DOMAIN 970 1036 ACYL CARRIER (ACP) 1.
 FT DOMAIN 2015 2082 ACYL CARRIER (ACP) 1.
 FT DOMAIN 3043 3109 ACYL CARRIER (ACP) 3.
 FT BINDING 999 999 PHOSPHOPANTHETHEINE (POTENTIAL).
 FT BINDING 2045 2045 PHOSPHOPANTHETHEINE (POTENTIAL).
 FT BINDING 3073 3073 PHOSPHOPANTHETHEINE (POTENTIAL).

FT CONFLICT 33 33 S -> F (IN REF. 1).
 FT CONFLICT 42 42 A -> G (IN REF. 1).
 FT CONFLICT 110 110 Q -> D (IN REF. 1).
 FT CONFLICT 113 115 RQA -> AQC (IN REF. 1).
 FT CONFLICT 139 139 L -> V (IN REF. 1).
 FT CONFLICT 259 259 L -> W (IN REF. 1).
 FT CONFLICT 309 309 R -> A (IN REF. 1).
 FT CONFLICT 478 480 TPA -> SRP (IN REF. 1).
 FT CONFLICT 596 596 MISSING (IN REF. 4).
 FT CONFLICT 648 648 A -> R (IN REF. 1).
 FT CONFLICT 680 682 RHV -> ETL (IN REF. 1).
 FT CONFLICT 694 698 EQSIT -> DKRIS (IN REF. 4).
 FT CONFLICT 788 788 M -> L (IN REF. 4).
 FT CONFLICT 939 940 PL -> LV (IN REF. 1).
 FT CONFLICT 1038 1038 N -> I (IN REF. 1).
 FT CONFLICT 1133 1133 H -> Q (IN REF. 1).
 FT CONFLICT 1310 1310 V -> C (IN REF. 1).
 FT CONFLICT 1333 1333 G -> V (IN REF. 1).
 FT CONFLICT 1384 1384 P -> R (IN REF. 1).
 FT CONFLICT 1582 1582 G -> E (IN REF. 1).
 FT CONFLICT 1677 1682 KRRADG -> E (IN REF. 1).
 FT CONFLICT 1700 1700 C -> S (IN REF. 1).
 FT CONFLICT 1755 1755 F -> K (IN REF. 1).
 FT CONFLICT 1787 1787 T -> S (IN REF. 1).
 FT CONFLICT 1801 1822 GAATGRVDLYEPDAFAKRPITIG ->
 FT CONFLICT 1915 1916 APSPGGLICMRCICETPDNR (IN REF. 1).
 FT CONFLICT 2075 2075 LG -> PK (IN REF. 1).
 FT CONFLICT 2079 2079 R -> C (IN REF. 1).
 FT CONFLICT 2141 2145 A -> V (IN REF. 1).
 FT CONFLICT 2445 2445 ARLTP -> LRSLN (IN REF. 1).
 FT CONFLICT 2485 2485 E -> Q (IN REF. 1).
 FT CONFLICT 2546 2567 ATDLF -> RQCS (IN REF. 1).
 FT CONFLICT 2608 2615 TVHQLFEETVQRHKDPVAVTN -> DGCISYKRLSSATK
 FT CONFLICT 2644 2645 TARLSHTN (IN REF. 1).
 FT CONFLICT 2713 2713 MSAAVLGV -> KCPPRCSAS (IN REF. 1).
 FT CONFLICT 2723 2723 KL -> NV (IN REF. 1).
 FT CONFLICT 2876 2880 D -> H (IN REF. 1).
 FT CONFLICT 2899 2900 GELCVA -> RALRG (IN REF. 1).
 FT CONFLICT 2958 2960 RF -> L (IN REF. 1).
 FT CONFLICT 2964 2964 EDR -> QDA (IN REF. 1).
 FT CONFLICT 3587 3587 R -> A (IN REF. 1).
 SQ SEQUENCE 401248 MW; 6B0B05A9FF32054D CRC64;
 Query Match 8.5%; Score 287.5; DB 1; Length 3587;
 Best Local Similarity 21.5%; Pred. No. 1e-13;
 Matches 131; Conservative 96; Mismatches 249; Indels 133; Gaps 19;
 QY 42 RELRIVCKTARRDLFGLSLVLRVLRRLRHQRAGHTIP-----RIQAVVQVQPERL 93
 DB 2503 RLKXAVADNPQPLSSLLLVTEKQALLKAWKGAIPVPTDKTVHQLFEETVQRHKDRP 2562
 QY 94 ALVDAGTCGECWTFQAQLDAYSNVANLFRQLGFAPGDVWAIFFLRGPRFVGLWGLAKAGM 153
 DB 2563 AV--TYNGQSWTYGELNANRLARILMDGISPDDRVRGVLTKPSLEMSAAVLGVKACA 2620
 QY 154 EALLNVNLRREPLAFCLGTSGAKALIFGEMVAVAEVSGLKSLKFCSDGLDPEGI 213
 DB 2621 AFVPIDPDYDQRIEYLQDSGAKLL-----LKOEGI 2652
 QY 214 -LPDTHLLDPLLKEASTAPLAQIPSKGMD-----RLFYITSTTGLPKAAIV 261
 DB 2653 SVFDSYTGVDILLGSKTILS-LPDENDEGNPETAVTAENLAMYITSTGTPQPKGVN 2711
 QY 262 VHSRYRMAAFGHAYRMOAAD-----VLYDCLPLYHSAGNIIGVQCQLIYGL 309
 DB 2712 EDHALVNLCLFWDHDAFSMTAEDRSARYAGGFDASIWENFTWTTCALHVDIDEALRDI 2771
 QY 310 TVVLRKFKFSRWDCCIKNYTVVOYICEICRYLLKQPVREAER-----RHRVRLAV 362
 DB 2772 VRL-----NDYFTNGVTITFL-----PTLAEQFMELENTSRLVLLTG 2810

